

Bildungsbedarfsanalyse auf Grundlage von Stellenanzeigen – Potenziale des Text Mining für das Lern-Service-Engineering

Frank Bensberg

Veröffentlicht in:
Multikonferenz Wirtschaftsinformatik 2012
Tagungsband der MKWI 2012
Hrsg.: Dirk Christian Mattfeld; Susanne Robra-Bissantz



Braunschweig: Institut für Wirtschaftsinformatik, 2012

Bildungsbedarfsanalyse auf Grundlage von Stellenanzeigen – Potenziale des Text Mining für das Lern-Service-Engineering

Frank Bensberg

Hochschule für Telekommunikation Leipzig, Department für Wirtschaft, 04277 Leipzig,
E-Mail: bensberg@hft-leipzig.de

Abstract

Die extrem kurzen Technologiezyklen in der ICT-Industrie stellen Unternehmen vor das Problem, Mitarbeiter zeit- und themenadäquat weiter zu qualifizieren. Für Bildungsanbieter erwächst damit die Herausforderung, relevante Bildungsthemen möglichst frühzeitig zu identifizieren, ökonomisch zu bewerten und ausgewählte Themen in Form geeigneter Leistungsangebote zur Marktreife zu bringen. Zur Entscheidungsunterstützung der Identifikationsphase wird in diesem Beitrag das bedarfsanalytische Instrument der Stellenanzeigenauswertung aufgegriffen und dessen Umsetzung mithilfe von Methoden des Text Mining exemplarisch demonstriert. Abschließend wird die Integration dieses Instruments mit dem Lern-Service-Engineering thematisiert und weiterführender Forschungsbedarf identifiziert.

1 Problemstellung

Der anhaltende Innovations- und Transformationsdruck im Bildungssektor hat dazu geführt, dass bei der Entwicklung von Bildungsangeboten in zunehmendem Maße eine *marktorientierte Sichtweise* eingenommen wird, die am Bildungsbedarf und den beruflichen Anforderungen der Zielgruppen ansetzt (vgl. [11], S. 22 f.). Diese marktorientierte Sichtweise wird auch im Rahmen des *Lern-Service-Engineerings* thematisiert, das die interdisziplinäre Bereitstellung und Verwendung von Prinzipien, Methoden und Werkzeugen für die zielorientierte Gestaltung von Lehr-/Lern-Leistungsangeboten intendiert (vgl. [5], S. 5).

Informatorische Grundlage für das Design solcher Leistungsangebote sind Planungsdaten, die z. B. Auskunft über marktseitig nachgefragte Bildungsthemen, Kompetenzen und Lernformate liefern. In der Wertschöpfungskette von Bildungsinstitutionen wird die systematische Gewinnung dieser Datengrundlage in der Aktivität der *Bildungs- bzw. Lernbedarfsermittlung* abgebildet (vgl. [16], S. 107 u. [11], S. 28). Eine besondere Herausforderung an die Lernbedarfsermittlung stellen solche Leistungsangebote, die Bildungsthemen aus der ICT-Industrie adressieren.

Stetig kürzer werdende Technologiezyklen und die hiermit einhergehende Wissenserrosion erfordern eine fundierte und vorausschauende Planung von Bildungsangeboten, die die nachfrageseitigen Anforderungen berücksichtigt.

Zur Gewinnung einer adäquaten Planungsgrundlage für Bildungsangebote werden in der Literatur unterschiedliche Verfahren diskutiert (vgl. [17], S. 331 u. [11], S. 29-32). Als eine potenziell adäquate Methode wird hier die *Stellenanzeigenanalyse* aufgegriffen, die auf Grundlage von Stellenanzeigen zeitnah Aussagen über berufliche, personenbezogene und sozial-kommunikative Anforderungen von Institutionen an spezifische Bewerbergruppen liefern kann (vgl. [15], S. 39 u. [7], S. 236 f.). Diese Methode ist in der Lage, den Qualifikationsbedarf unterschiedlicher Berufsfelder abzubilden und inhaltliche Verschiebungen im Zeitablauf zu verdeutlichen (vgl. [15], S. 40). Neben diesen materiellen Eigenschaften ist die Methode auch aus forschungsökonomischer Perspektive attraktiv, da die empirische Basis in der Regel öffentlich zugänglich ist und – z. B. über Jobportale – in Form elektronischer Dokumente vorliegt.

Zur Analyse elektronischer Dokumente werden in der Wirtschaftsinformatik Methoden des *Text Mining* thematisiert (vgl. [14]). Diese gestatten die Auswertung unstrukturierter Datenbestände und stehen mittlerweile in Form marktgängiger Analysesysteme zur Verfügung. Im Rahmen dieses Beitrags sollen die informatorischen Potenziale solcher Text Mining-Systeme zur Stellenanzeigenanalyse transparent gemacht werden. Zu diesem Zweck wird zunächst ein kurzer Überblick über die technischen Grundlagen des Text Mining gegeben. Hierauf aufbauend werden das prototypische Design einer Text Mining-Lösung für die Stellenanzeigenanalyse dargestellt und Ergebnisse präsentiert, die im Rahmen der Projektstudie JobMining@HfTL an der Hochschule für Telekommunikation (HfTL) gewonnen werden konnten. Abschließend wird die Integration mit dem Konzept des Lern-Service-Engineerings aufgezeigt und weiterführender Forschungsbedarf identifiziert.

2 Technische Grundlagen des Text Mining

Unter dem Begriff des Text Mining ist ein weitgehend automatisierter, analytischer Prozess zu verstehen, der zur Gewinnung von neuem und potenziell nützlichem Wissen aus Textdokumenten dient (vgl. [10], S. 287). Im Gegensatz zum Data Mining, das zur Analyse *strukturierter Daten* eingesetzt wird, stellt sich beim Text Mining die Herausforderung, mithilfe statistischer und linguistischer Analysemethoden weitgehend *unstrukturierte Daten* zu verarbeiten. Diese liegen typischerweise in Form elektronisch gespeicherter Dokumente (z. B. HTML- oder XML-Dateien) vor, deren Inhalte nur relativ schwach ausgeprägte Beziehungen zueinander aufweisen. Die hierfür erforderlichen Verarbeitungsschritte sind nun anhand eines generischen Architekturmodells für Text Mining-Systeme zu erläutern (vgl. Bild 1).

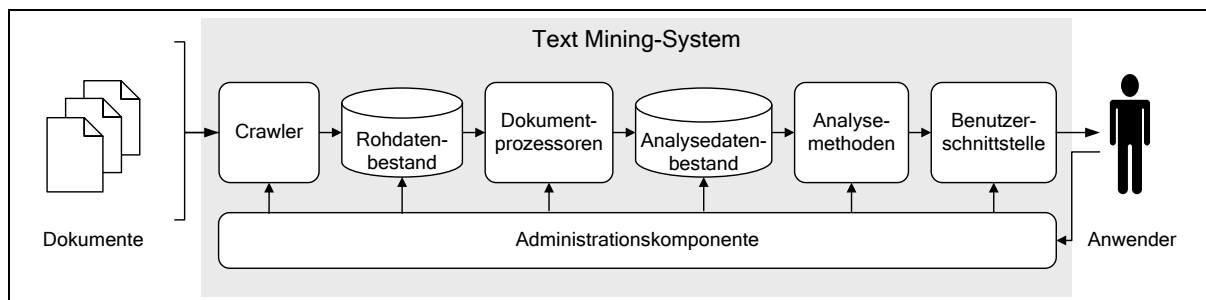


Bild 1: Architekturmodell eines Text Mining-Systems (in Anlehnung an [4], S. 15)

Grundlage des Text Mining-Prozesses sind Dokumente, die entweder als Online-Quellen über Internet-Dienste bereitgestellt werden (z. B. HTML-Seiten, E-Mail-Dokumente), oder aber in Dateiform auf einem lokalen Speichermedium vorliegen (Offline-Quellen). In einem ersten Schritt sind analyserelevante Dokumente zu erfassen, sodass diese in einem lokalen *Rohdatenbestand* im Text Mining-System vorliegen und für die Weiterverarbeitung zur Verfügung stehen. Diese Aufgabe wird in dem dargestellten Architekturmodell von einer *Crawlerkomponente* übernommen, die Dokumente aus Datenquellen extrahiert.

Der somit generierte Rohdatenbestand wird anschließend mithilfe von *Dokumentprozessoren* verarbeitet, die der Extraktion relevanter Merkmale aus den gesammelten Dokumenten dienen (*Information Extraction*, vgl. [2], S. 45 f.). Diese Prozessoren ermitteln z. B. die Quellsprache des jeweiligen Dokuments, zerlegen Wortkomposita in Einzelelemente und extrahieren aussagekräftige Wortarten bzw. Wortfolgen. Um die Dokumente für Such- und Filteroperationen vorzubereiten, wird i.d.R. auch ein Index aufgebaut, der jedem Dokument eine Reihe von Schlüsselwörtern zuordnet.

Nach der Merkmalsextraktion liegt ein *Analysedatenbestand* vor, auf den *Methoden* zur Textanalyse angewendet werden können. Geläufig sind hierzu etwa Verfahren zur Häufigkeitsanalyse von Worten bzw. Wortarten, zur Identifikation von Assoziationen zwischen Begriffen sowie zur Segmentierung und Klassifikation von Dokumenten (vgl. [10], [4], [3]). Zur Parametrisierung der Analysemethoden und zur Visualisierung der Ergebnisse steht dem Anwender eine *Benutzerschnittstelle* als Front-End zur Verfügung. Die in Bild 1 dargestellte *Administrationskomponente* dient hingegen zur Konfiguration und Überwachung einzelner Architekturelemente. So ist etwa für die Crawlerkomponente festzulegen, ob die Datenquellen kontinuierlich oder anhand eines benutzerdefinierten Zeitplans zu durchsuchen sind, während für die Dokumentprozessoren zu konfigurieren ist, wann der Index reorganisiert werden soll.

Das dargestellte Architekturmodell wird im Folgenden als Bezugsrahmen zur Gestaltung einer konkreten Text Mining-Lösung zur Analyse von Stellenanzeigen herangezogen.

3 Design einer Text Mining-Lösung zur Stellenanzeigenanalyse – das Projekt JobMining@HfTL

3.1 Fachliche Problemstellung des Projekts JobMining@HfTL

Die Hochschule für Telekommunikation Leipzig (HfTL) ist eine private Fachhochschule, die sich in Trägerschaft der Deutschen Telekom AG befindet und technisch orientierte Studiengänge mit Fokus auf den Informations- und Telekommunikationssektor anbietet. Dieses akademische Ausbildungsportfolio wird zudem durch Weiterbildungsangebote ergänzt, die zur Qualifizierung von Arbeitnehmern in eng abgrenzbaren Themenfeldern dienen. Die Studiengänge werden in unterschiedlichen Formen angeboten: neben dem klassischen Präsenzstudium werden auch berufsbegleitende und duale Studiengänge offeriert (vgl. [12]). Derzeit verfügt die HfTL über mehr als 700 Studierende, wobei der Anteil derjenigen Studierenden, die an berufsbegleitenden sowie dualen Studienformen partizipieren und somit in erheblichem Umfang auf E-Learning-Szenarien zurückgreifen, mittlerweile dominiert.

Aufgrund der hohen Entwicklungsdynamik der Qualifikationsanforderungen an Arbeitnehmer im ICT-Sektor stellt sich der HfTL die Herausforderung, relevante Bildungsthemen möglichst frühzeitig zu identifizieren und diese konstruktiv in die Entwicklung der Bildungsangebote einfließen

zu lassen. Zu diesem Zweck ist im Jahr 2011 das Projekt JobMining@HfTL initiiert worden, das den Aufbau eines analytischen Informationssystems zur Auswertung von Stellenanzeigen zum Gegenstand hat. Fragestellungen, die den ausschlaggebenden Impuls für dieses Projekts geliefert haben, werden im Folgenden kurz skizziert:

- Wie verteilen sich die geforderten *akademischen Abschlüsse* auf Stellenanzeigen?
- Welche *fachlichen Ausbildungsinhalte* werden im Rahmen der Stellenanzeigen als Anforderungen an potenzielle Bewerber formuliert?
- Können Aussagen über die *zeitliche Entwicklung* von geforderten Ausbildungsinhalten getroffen werden, um qualifikationsbezogene Trends möglichst frühzeitig zu erkennen?

Um einen Beitrag zur Deckung des skizzierten Informationsbedarfs zu liefern, sind im Rahmen des Projekts JobMining@HfTL Stellenanzeigen der Deutschen Telekom AG (DTAG) als eines der weltweit führenden Dienstleistungsunternehmen des ICT-Sektors untersucht worden. Im Folgenden wird – aufbauend auf der Darstellung der Datengrundlage – erörtert, wie die oben angeführten Fragestellungen mithilfe von Analysemethoden des Text Mining beantwortet werden können.

3.2 Datengrundlage

Als Ausgangspunkt für die Text Mining-Analyse wurden die von der DTAG publizierten Stellenanzeigen selektiert, die im Internet als Online-Quellen über ein Jobportal unter dem URL <http://www.telekom.com/your-chance> vorgehalten werden. Dieses Jobportal informiert interessierte Bewerber über Vakanzen und stellt diese stellenbezogenen Informationen personalisiert zur Verfügung. Hierfür definieren Interessenten zuvor einen *Jobagenten* (vgl. [18], S. 206), der dafür sorgt, dass Bewerber per E-Mail wöchentlich über neue Ausschreibungen informiert werden. Die übermittelte E-Mail (*Jobmail*) enthält dabei einen Verweis auf eine Liste mit entsprechenden Stellenanzeigen auf der Website des Jobportals. Die einzelnen Segmente einer Stellenanzeige werden in Bild 2 exemplarisch dargestellt.

Das skizzierte Beispiel enthält eine Stellenanzeige für einen *IT-Service Manager* (① Stellenbezeichnung), die zunächst eine Reihe von Detailangaben (②) zur Stelle auflistet (z. B. *Land*, *Region* und *Bewerbungsschluss*). Im Anschluss folgen die Vorstellung des ausschreibenden Unternehmensbereichs (③), die eigentliche Stellenbeschreibung mit Aufzählung der Aufgabfelder (④), die erwünschten Bewerberqualifikationen (⑤) mit den spezifischen Fähigkeiten, Kenntnissen und Verhaltensweisen, sowie ein abschließender Text mit stellenbezogenen Karriereinformationen (⑥), die weitere Anreize zur Bewerbung kommunizieren sollen (vgl. [15], S. 127 f.).

Um diese Felder extrahieren zu können, ist mithilfe des Web Application Frameworks *Ruby on Rails* ein *Screen Scraper* (vgl. [9]) entwickelt worden, der das gezielte Auslesen von Stellenanzeigen aus der Präsentationsschicht des Jobportals gestattet. Mithilfe dieses Screen Scrapers wird – sobald eine neue Jobmail verfügbar ist – der per E-Mail zugestellte Verweis auf die Stellenanzeigenliste des Jobportals aktiviert und sämtliche Anzeigen der Liste ausgelesen. Im Zuge des Projekts JobMining@HfTL sind im Zeitraum von Juni bis August 2011 insgesamt 1.108 Stellenanzeigen extrahiert worden, die die Grundlage für die Anwendung von Analysemethoden des Text Mining bilden.

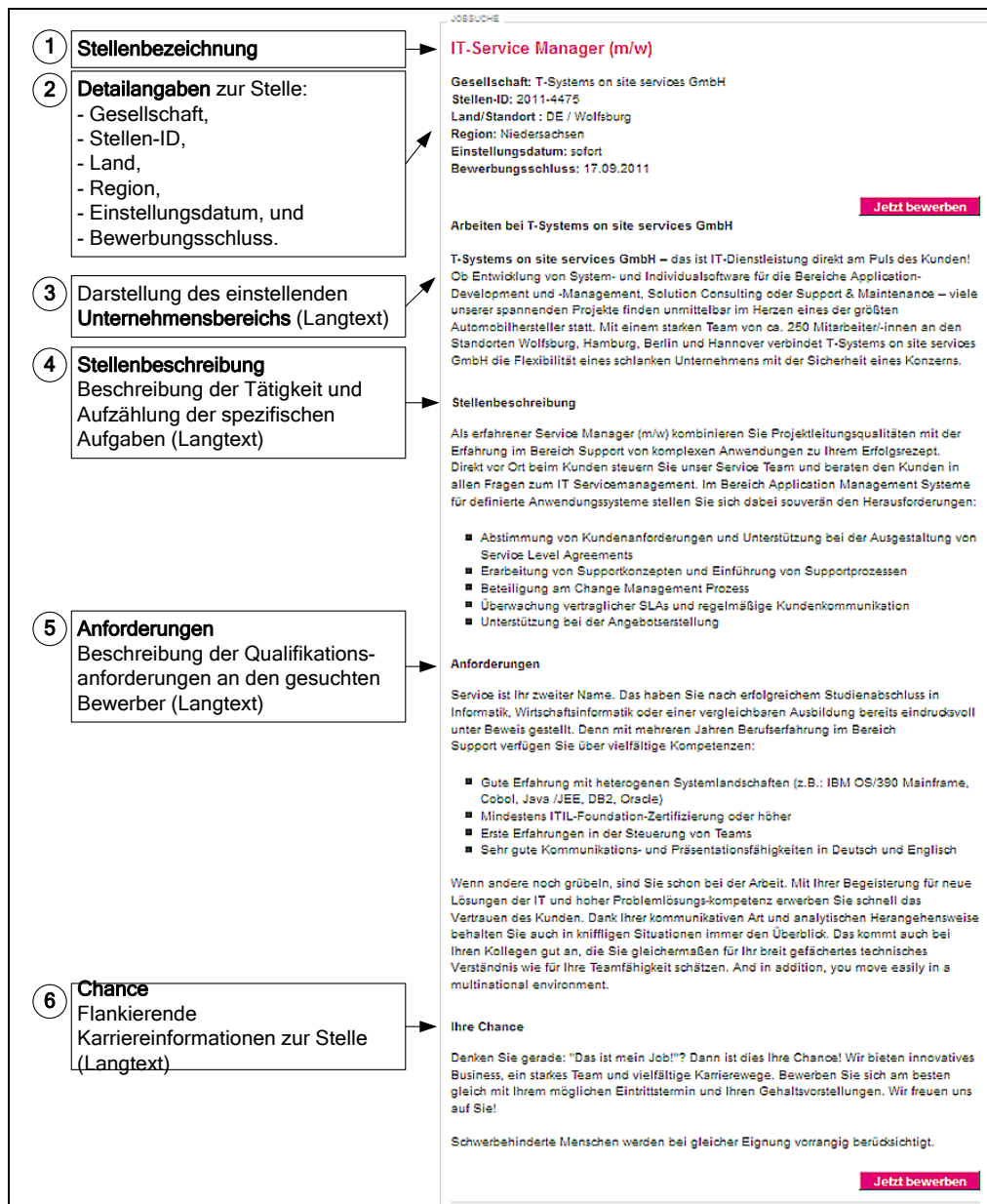


Bild 2: Aufbau einer Stellenanzeige im Jobportal der Deutsche Telekom AG

Zur Analyse ist das Softwaresystem *IBM Content Analytics* eingesetzt worden, dessen *Crawler-komponente* und *Dokumentprozessoren* (vgl. Bild 1) zur Verarbeitung der Dokumentkollektion konfiguriert wurden. Die einzelnen Konfigurationsparameter werden im Rahmen dieses Beitrags nicht vertieft; statt dessen wird auf die produktspezifische Darstellung in [20] verwiesen. Im Folgenden wird die Anwendung ausgewählter Analysemethoden zur Beantwortung der in Abschnitt 3.1 identifizierten Fragestellungen dargestellt.

3.3 Analyse der akademischen Abschlüsse

Um die Verteilung der geforderten akademischen Abschlüsse transparent zu machen, sind Analysetechniken des *Information Retrieval* eingesetzt worden. Unter diesem Begriff werden Techniken verstanden, die die Suche nach Dokumenten unterstützen, die inhaltlich zu einer Suchanfrage eines Informationsnachfragers passen. Finden sich die Inhalte der Suchanfrage in einem Dokument, werden diese als Treffer ausgegeben (vgl. [2], S. 43-46).

Im Zuge der Analyse sind die Stellenanzeigen nach Schlüsselwörtern durchsucht worden, die für typische ingenieur- bzw. wirtschaftswissenschaftliche Studiengänge stehen (z. B. *Nachrichten-* und *Elektrotechnik*, *Informatik*, *Wirtschaftsinformatik*, *Betriebswirtschaft*). Dies wird vom eingesetzten Text Mining-System durch eine Abfragesprache unterstützt, mit der komplexe Suchausdrücke formuliert werden können. Bild 3 zeigt das Ergebnis der Volltextsuche in der Dokumentkollektion nach alternativen Begriffen für den Studiengang *Wirtschaftsinformatik*. Nach Eingabe des Suchausdrucks werden die gefundenen Stellenanzeigen in einer Ergebnisliste dargestellt und stehen zur Detailanalyse zur Verfügung.

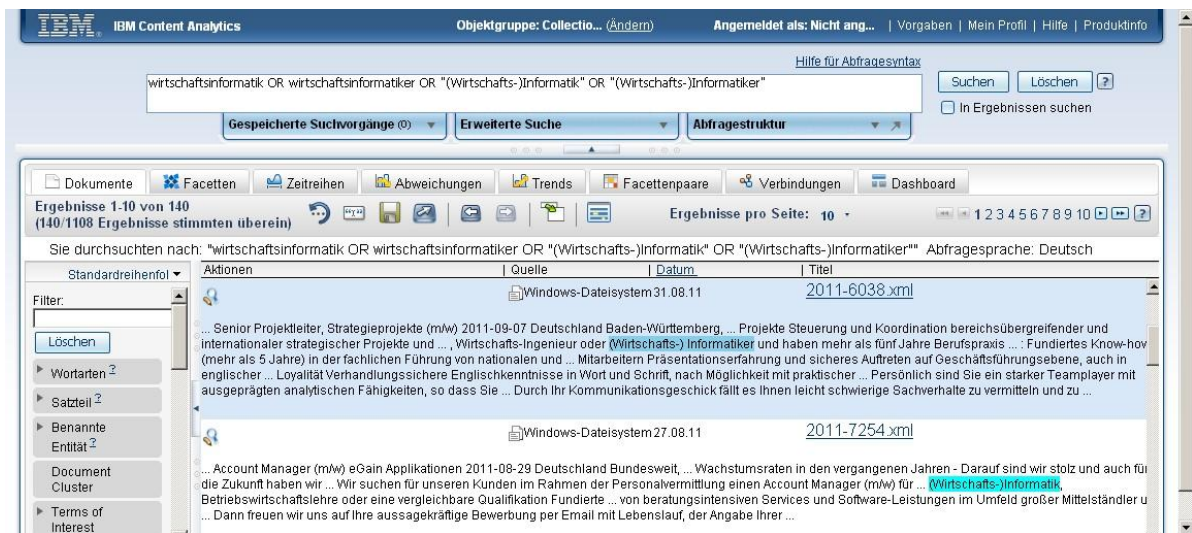


Bild 3: Volltextsuche in den Stellenanzeigen mithilfe von Schlüsselwörtern

Insgesamt konnten 140 Stellenanzeigen identifiziert werden, in denen explizit auf den Studienabschluss *Wirtschaftsinformatik* Bezug genommen wird (13% sämtlicher Stellenanzeigen). Dabei konnte auch festgestellt werden, dass in dem Analysezeitraum häufiger Absolventen der *Informatik* (27%) und der *Betriebswirtschaftslehre* (24%) gesucht wurden, während indes *nachrichten-* bzw. *elektrotechnische* Abschlüsse nur in etwa 7% der Stellenanzeigen referenziert werden.

Mit diesen Analyseergebnissen wird zunächst die Frage nach der *horizontalen Konkurrenz-situation* zwischen unterschiedlichen Studiengängen beantwortet (vgl. [15], S. 163). Diese quantitativen Informationen können für Bildungsanbieter hilfreich sein, um Entscheidungen über die *Neueinführung* von Studiengängen zu fundieren oder aber Impulse zur *Revision* bestehender Studiengänge zu generieren (vgl. [11], S. 6).

3.4 Analyse der fachlichen Ausbildungsinhalte

Zur Analyse der Ausbildungsinhalte sind die Stellenanzeigen auf die Nennungen einzelner *Fachbegriffe* untersucht worden, die somit die grundlegende Analyseeinheit darstellen (vgl. [11], S. 120 f.) und bei Stellenanzeigen insbesondere in der Beschreibung der Qualifikationsanforderungen (s. Bild 2) auftreten. Diese Aufgabenstellung wird von dem Text Mining-System durch eine *grammatikalische Analyse* unterstützt, die aus den Stellenanzeigen automatisch Substantive, Verben, Adjektive und andere Wortarten extrahiert. In der Terminologie des eingesetzten Analysesystems bilden diese unterschiedliche Wortarten sogenannte *Facetten*, die unterschiedliche Sichten (Analysedimensionen) auf die Dokumentkollektion erschließen.

Als Analysefacette für die fachlichen Ausbildungsinhalte sind die *Substantive* herangezogen worden, die vom Text Mining-System listenorientiert dargestellt und mit Häufigkeiten ausgezeichnet werden (s. Bild 4). Die Häufigkeitsangabe bezieht sich dabei auf die Anzahl der Stellenanzeigen, in denen das Schlüsselwort aufzufinden ist. So kann der dargestellten Liste z. B. entnommen werden, dass das Schlüsselwort *Excel* in 85 Stellenanzeigen auftritt.

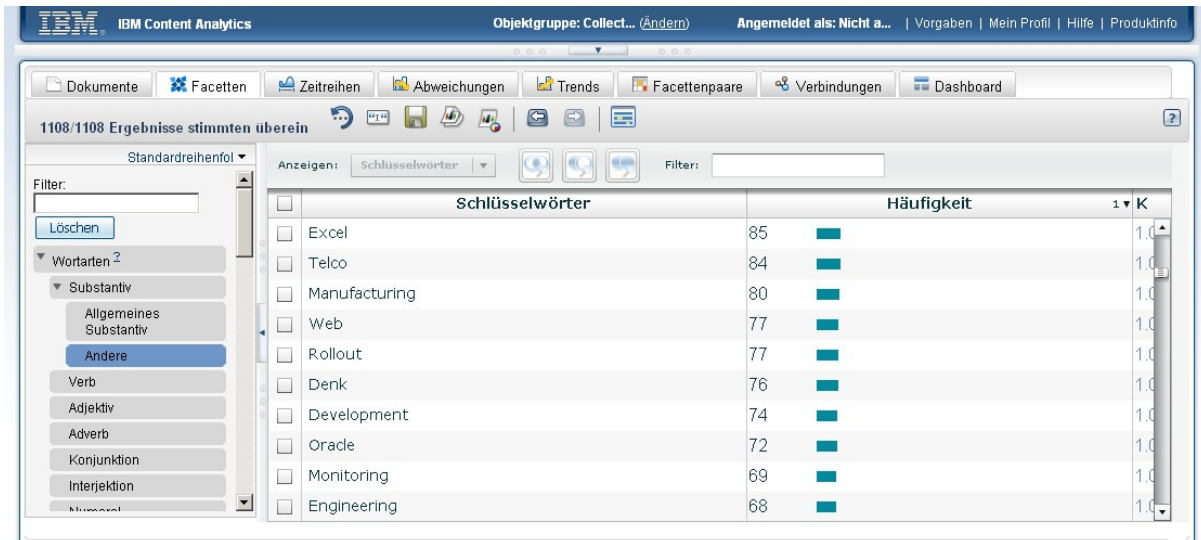


Bild 4: Listenorientierte Darstellung in der Analysefacette Substantive

Bei informatikorientierten Studiengängen befinden sich Ausbildungsinhalte im Spannungsfeld zwischen dem längerfristig gültigen *Konzeptwissen* und dem tendenziell eher kurzfristig ausgerichteten *Produktwissen* (vgl. [8], S. 23 f.). Um diesem Sachverhalt Rechnung zu tragen, sind auf Grundlage der identifizierten Schlüsselwörter induktiv drei unterschiedliche *Kategorien* gebildet worden (vgl. [13], S. 114 ff.), die unterschiedliche Gruppen von Ausbildungsthemen reflektieren:

- In der Kategorie *Konzepte & Methoden* werden solche Ausbildungsthemen angesiedelt, die das längerfristige Konzeptwissen aus informatikorientierten Fachdisziplinen konstituieren (z. B. serviceorientierte Architekturen, SOA).
- Mit der Kategorie *Hersteller & Produkte* werden Nennungen einzelner Produzenten, Marken und Produktbezeichnungen im ICT-Sektor erfasst (z. B. Oracle, Eclipse). Diese Kategorie bildet das geforderte Produktwissen ab.
- Die Kategorie *Sprachen* umfasst Programmiersprachen (z. B. Java), Datenabfragesprachen (z. B. SQL), Modellierungssprachen (z. B. UML) und Auszeichnungssprachen (z. B. HTML). Dabei werden Sprachen als vermittelndes *Bindeglied* zwischen dem Konzept- und dem Produktwissen betrachtet.

Auf Grundlage der gegebenen Häufigkeiten sind die führenden zehn Schlüsselwörter (Top 10) für jede Kategorie ermittelt worden. Diese Ausbildungsthemen werden in der folgenden Abbildung mit ihrer relativen Häufigkeit dargestellt und im Folgenden kurz erörtert.

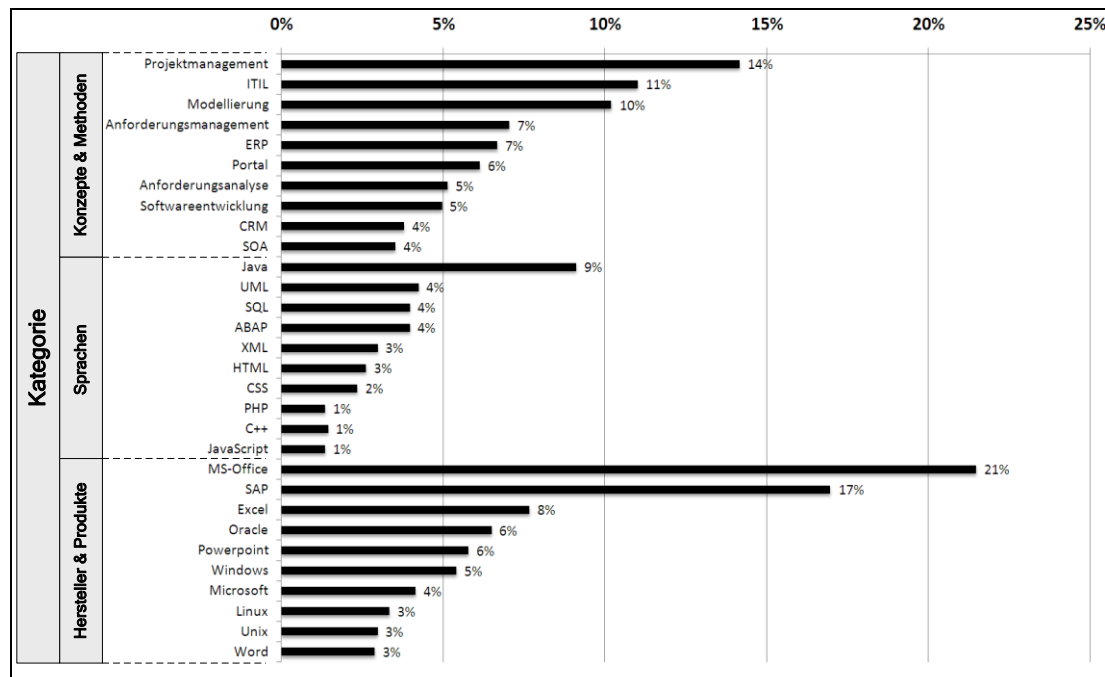


Bild 5: Darstellung der häufigsten Ausbildungsthemen nach Kategorien (Top 10)

- In der Kategorie *Konzepte & Methoden* werden von potenziellen Bewerbern insbesondere Kompetenzen in den Bereichen des *Projektmanagements*, des IT-Managements mit *ITIL* sowie der *Modellierung* abgerufen. In Anbetracht des anhaltenden Transformationsprozesses in der Telekommunikationsindustrie und dem hiermit einhergehenden Restrukturierungsbedarf sind diese Qualifikationsanforderungen als plausibel nachvollziehbar zu bewerten (vgl. [6]).
- Die Kategorie *Sprachen* zeigt deutlich, dass als dominierende Programmiersprachenkompetenz *Java* abgerufen wird. Dieses Ergebnis deckt sich auch mit dem aktuellen TIOBE-Index, in dem Java als beliebteste Programmiersprache geführt wird (vgl. [19]). An zweiter Stelle in dieser Kategorie steht die *Unified Modeling Language* (UML) als Modellierungssprache. Als interessant wird dabei erachtet, dass anderen Notationen – insbesondere der neueren *Business Process Model and Notation* (BPMN 2.0) und *ARIS* – in den Stellenanzeigen nur geringfügige Bedeutung beigemessen wird.
- Die letzte Kategorie *Hersteller & Produkte* verdeutlicht die Relevanz von Qualifikationen im Umgang mit Bürosoftware und konkreten ERP-Produkten. So verweisen nicht nur 21% der Stellenanzeigen auf Kenntnisse über das Softwarepaket *MS-Office*, vielmehr werden explizit auch die Einzellösungen *Excel*, *Powerpoint* und *Word* angeführt. Außerdem werden – nahezu erwartungsgemäß – in erheblichem Ausmaß Qualifikationen zu Lösungen der Marke *SAP* abgerufen.

Als neues und potenziell nützliches Wissen des Text Mining werden insbesondere diejenigen Ergebnisse angesehen, die durch die Analyse der Kategorien *Konzepte & Methoden* sowie *Sprachen* gewonnen werden konnten. Mithilfe dieser Ergebnisse ist im Rahmen einer Gap-Analyse zu prüfen, ob auf der Nachfrageseite dominierende Fachthemen existieren, die von den aktuellen Bildungsangeboten nicht oder nur unzureichend abgedeckt werden. Um eventuell bestehende Lücken auszugleichen, können z. B. Entscheidungen zur nachfrageorientierten Adjustierung der Bildungsinhalte entsprechender Studiengänge oder Weiterbildungsangebote gefällt werden.

3.5 Analyse der zeitlichen Entwicklung von Bildungsthemen

Die zuvor erörterten Ergebnisse weisen die Problematik auf, dass sie sich auf sämtliche Stellenzeigen der untersuchten Datengrundlage beziehen. Infolgedessen können keine Aussagen über die zeitliche Entwicklung einzelner Bildungsthemen getroffen werden. Da jedoch gerade im ICT-Sektor mit einer hohen Dynamik bei der Adoption neuer Konzepte und Technologien zu rechnen ist, ist die zuvor dargestellte Analyse um die zeitliche Dimension zu ergänzen, damit themenbezogene Trends frühzeitig erkannt werden können.

Das eingesetzte Text Mining-System bietet hierzu eine Sicht zur *Trendanalyse* an, mit der die Häufigkeiten von Schlüsselwörtern im Zeitablauf visualisiert und starke Häufigkeitszunahmen gesondert signalisiert werden. Diese Funktionalität setzt allerdings voraus, dass ein *zeitliches Ordnungsmerkmal* in der Datengrundlage existiert. Im Hinblick auf die in Bild 2 dargestellte Datengrundlage wurde zu diesem Zweck das Merkmal *Bewerbungsschluss* aus den Detailangaben der Stellenanzeigen selektiert, da dieses durchgängig über eine gültige Datumsangabe verfügt. Als Beispiel für die Trendvisualisierung wird in der folgenden Abbildung die zeitliche Entwicklung des Schlüsselworts *ITIL* in den Stellenanzeigen mit einer Auflösung auf Tagesebene dargestellt.

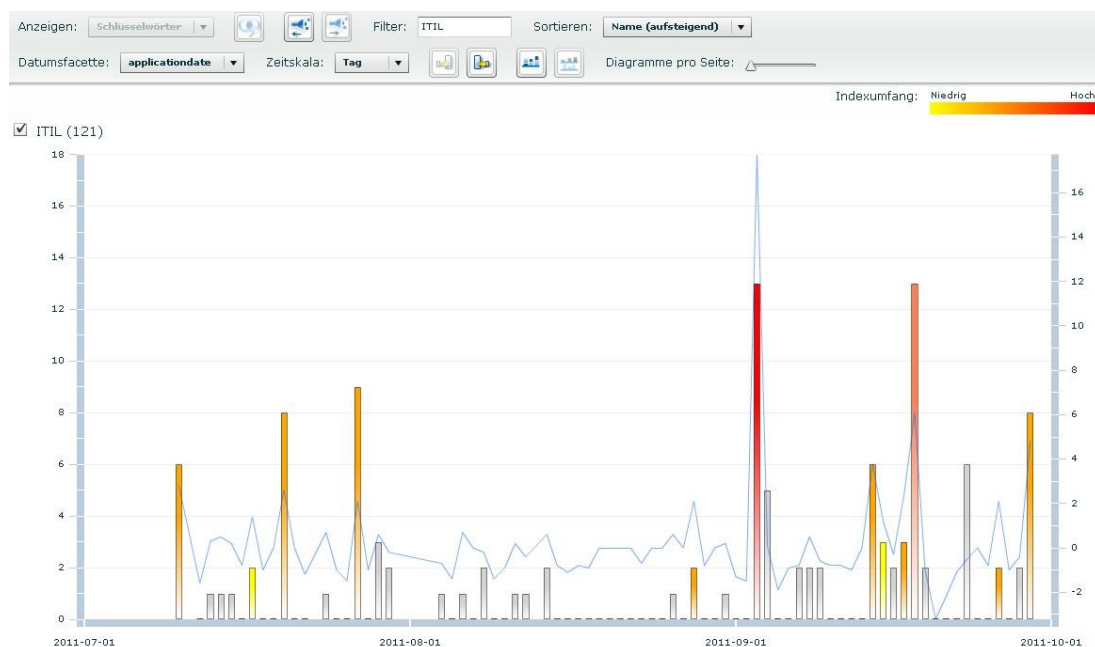


Bild 6: Zeitliche Entwicklung des Schlüsselworts ITIL

In dem *Balkendiagramm* wird die absolute Häufigkeit derjenigen Stellenanzeigen über die Zeitachse dargestellt, in denen das Schlüsselwort ITIL aufgetreten ist. So kann z. B. dem ersten Datenpunkt des Diagramms entnommen werden, dass sechs Stellenanzeigen (Skalierung auf der linken Ordinatenachse) mit dem Bewerbungsschluss 10.7.2011 das Fachthema *ITIL* enthalten. Darüber hinaus wird dieses Balkendiagramm durch ein *Liniendiagramm* ergänzt, das einen Indexwert für die Häufigkeitszunahme abbildet (Skalierung auf der rechten Ordinatenachse). Dieser *Zunahmeindex* drückt aus, in welchem Ausmaß die tatsächliche Häufigkeit von der erwarteten, durchschnittlichen Häufigkeit abweicht, die auf Grundlage der Ist-Werte der vergangenen Perioden ermittelt wird (vgl. [20], S. 236 f.). Ein hoher Zunahmeindexwert wird im Balkendiagramm zudem durch eine farbliche Codierung (Color Coding) des entsprechenden Balken in intensiver werdenden Farbtönen signalisiert.

Insgesamt ist der Zeitreihe zu entnehmen, dass an den meisten Tagen überhaupt keine oder nur sehr wenige Begriffsnennungen auftreten, die dann allerdings an einigen wenigen Tagen sprunghaft ansteigen. Dabei konzentrieren sich die Werte auf den letzten Analysemonat, der folglich ein Ballungsintervall bildet. Aufgrund des sporadischen Charakters der Zeitreihe und des kurzen Analysezeitraums ist die Anwendung traditioneller Prognosemodelle zur künftigen Bedarfsvorhersage problematisch. Allerdings verdeutlicht das skizzierte Beispiel gut, dass der Anwender mit der Trendanalyse in die Lage versetzt wird, Häufigkeitsveränderungen auf der Ebene einzelner Bildungsthemen visuell zu entdecken (Visual Data Mining) und daraufhin entsprechende Maßnahmen zu ergreifen. Anhaltende *Häufigkeitszunahmen* können z. B. zum Anlass genommen werden, um das eigene Ausbildungsportfolio in Bezug auf die Abdeckung des Themas kritisch zu prüfen. Andererseits können *Häufigkeitsabnahmen* ein Signal dafür sein, dass bestimmte Themen verfallen und dem Ende ihres Lebenszyklus entgegenschreiten.

Die vorangehende Darstellung der projektbezogenen Analyseergebnisse hat gezeigt, dass mithilfe des Text Mining interessante Informationen aus Stellenanzeigen extrahiert werden können, die für die fachlich-inhaltliche Ausgestaltung von Ausbildungsangeboten verwendbar sind. Infolgedessen ist nun die Integration dieses Instruments in das Konzept des Lern-Service-Engineerings herauszuarbeiten.

4 Integration mit dem Lern-Service-Engineering

Als Grundlage für die Integration des vorgestellten Ansatzes der Stellenanzeigenanalyse wird hier das Rahmenkonzept für die systematische Dienstleistungsentwicklung von Bullinger und Schreiner herangezogen, das ein idealtypisches Vorgehensmodell umfasst (vgl. [1], S. 73). Dieses Vorgehensmodell wird in Bild 7 für den Gegenstandsbereich des Lern-Service-Engineerings adaptiert dargestellt. Die Integration der Stellenanzeigenanalyse mithilfe des Text Mining erfolgt dabei in der zweiten Phase, die generell der *Anforderungsanalyse* von Lern-Services dient. Im Zuge dieser Phase sind Methoden einzusetzen, die der systematischen Erhebung von Kunden- bzw. Marktanforderungen dienen (vgl. [1], S. 74). Wie die in den vorangehenden Abschnitten skizzierten Ergebnisse zeigen, eignet sich die Stellenanzeigenanalyse zur Identifikation häufig nachgefragter Fachthemen von Institutionen bei der Bewerbersuche, sodass ein Beitrag zur Abschätzung des Marktpotenzials und der thematischen Relevanz von Lern-Services bereitgestellt werden kann.

Dabei ist allerdings hervorzuheben, dass die Stellenanzeigenanalyse lediglich *eine* Methode darstellt, um Anforderungen an Lern-Services zu erheben. In der Literatur wird eine Reihe weiterer Verfahren diskutiert, die generell dem Zweck der Bedarfsermittlung für Lehr- bzw. Lernangebote dienen. So werden neben *text-* bzw. *inhaltsanalytischen Methoden* insbesondere auch *kommunikationsorientierte Verfahren* diskutiert, wie z. B. Experteninterviews, Workshops, Gruppendiskussionen und Delphi-Studien (vgl. [11], S. 29-32). Folglich ist davon auszugehen, dass sich in der Praxis des Lern-Service-Engineerings ein Methodenmix als tragfähig erweisen wird, der unterschiedliche Grundformen der Datenerhebung und -auswertung kombiniert. Die im Rahmen der vorgestellten Projektstudie JobMining@HfTL gewonnenen Erkenntnisse deuten jedoch auch darauf hin, dass sich das Verfahren der Stellenanzeigenanalyse weitgehend automatisieren lässt, sodass langfristig Kostenvorteile gegenüber kommunikationsorientierten Verfahren zu erwarten sind.

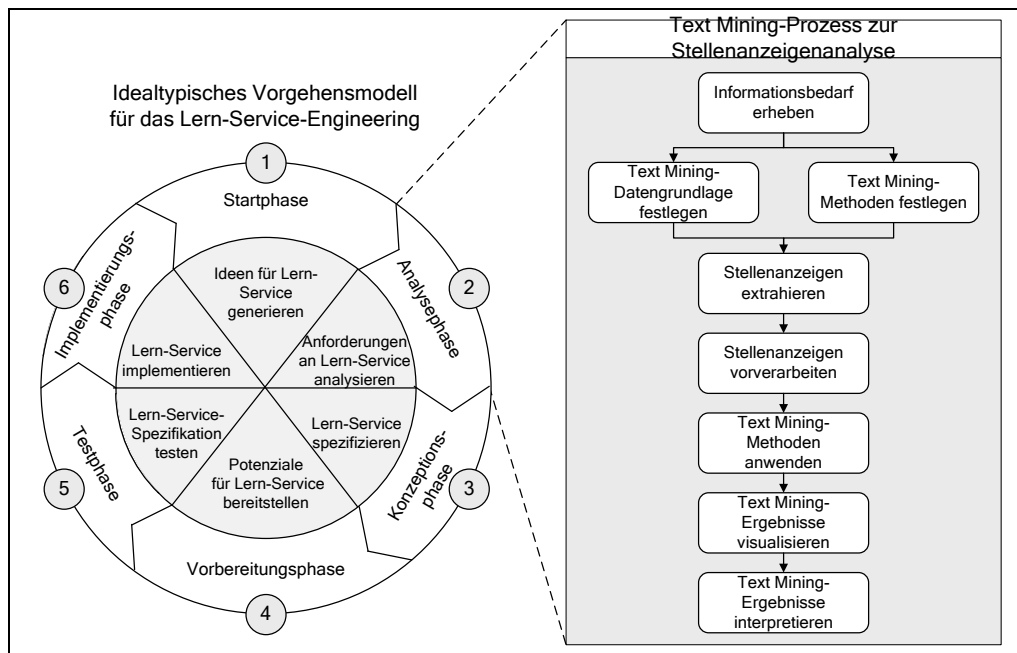


Bild 7: Integration des Text Mining in das Lern-Service-Engineering

Weiterführender Forschungsbedarf ergibt sich zunächst in Bezug auf die *Datengrundlage* und die *Methoden* des in Bild 7 cursorisch dargestellten Text Mining-Prozesses. So sind im Rahmen der Projektstudie nur die Stellenanzeigen eines einzigen Jobportals ausgelesen worden. Um eine breitere empirische Basis zu gewinnen, sind Stellenanzeigen aus weiteren Quellen für die Analyse verfügbar zu machen, wobei nahezu zwangsläufig Datenintegrationsprobleme zu erwarten sind. Außerdem stehen im Bereich des Text Mining weitere Analysemethoden zur Verfügung (z. B. zur Segmentierung und Klassifikation), die im Rahmen dieses Beitrags nicht erörtert worden sind. Im Zuge weiterer Projektstudien ist zu klären, ob diese Methoden für die Prozesse des Lern-Service-Engineerings weitere interessante Informationen liefern können.

Schließlich ist auch zu ermitteln, welchen Beitrag das vorgestellte Instrument für die zielorientierte Gestaltung von Lehr-/Lern-Leistungsangeboten liefert. Diese Fragestellung kann z. B. im Rahmen von Evaluationsstudien beantwortet werden, die sich mit den effektivitäts- und effizienzbezogenen Auswirkungen dieses Instruments auf die *Entscheidungsprozesse* im Umfeld des Lern-Service-Engineerings auseinandersetzen.

5 Literatur

- [1] Bullinger, H-J; Schreiner, P (2006): Service Engineering – Ein Rahmenkonzept für die systematische Entwicklung von Dienstleistungen. In: Bullinger, H-J, Scheer, A-W (Hrsg.), *Service Engineering*. 2. Auflage. Springer, Berlin.
- [2] Butscher, R (2005): Text Mining in der Konsumentenforschung unter besonderer Berücksichtigung von Produktontologien. Dissertation, Friedrich-Alexander-Universität Nürnberg-Erlangen.
- [3] Felden, C (2007): Text Mining als Anwendungsbereich von Business Intelligence. In: Chameni, P, Gluchowski, P (Hrsg.), *Analytische Informationssysteme*. Springer, Berlin.

- [4] Feldman, R; Sanger, J (2006): The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data. Cambridge University Press, New York.
- [5] Fink, C; Gabriel, R; Gersch, M; Lehr, C; Weber, P (2011): Lernservice-Engineering – Eine ökonomische Perspektive auf technologiegestütztes Lernen. In: Ebner, E, Schön, S (Hrsg.), *Lehrbuch für Lernen und Lehren mit Technologien*. BoD, Norderstedt.
- [6] Fuß, F (2010): Die Kunden, die Kultur und die Kontinuität – Über eine der größten Transformationen in der Telekommunikationsindustrie. In: Mohr, N, Büning, N, Hess, U, Fröbel A (Hrsg.), *Herausforderung Transformation*. Springer, Berlin.
- [7] Grob, H L; Lange, W (1996): Zum Wandel des Berufsbildes bei Wirtschaftsinformatikern: Eine empirische Analyse auf der Basis von Stellenanzeigen. WIRTSCHAFTSINFORMATIK 38(2):236-241.
- [8] Hartmann, W; Näf, M; Reichert, R (2007): Informatikunterricht planen und durchführen. Springer, Berlin.
- [9] Hemenway, K; Calishain, T (2003): Spidering Hacks – 100 Industrial Strength Tips & Tricks. O'Reilly, Sebastopol.
- [10] Hippner, H; Rentzmann, R (2006): Text Mining. Informatik Spektrum 29(4):287-290.
- [11] Hörmann, C (2007): Die Delphi-Methode in der Studiengangsentwicklung. Dissertation, Pädagogische Hochschule Weingarten.
- [12] Lecke, M (2010): Bologna@Telekom – Ein Beispiel für die Multioptionalität des Studierens. In: Keuper, F, Hogenschurz, B (Hrsg.), *Professionelles Sales & Service Management*. Gabler, Wiesbaden.
- [13] Mayring, P (2002): Qualitative Sozialforschung. 5. Auflage. Beltz, Weinheim.
- [14] Meier, M; Beckh, M (2000): Text Mining. WIRTSCHAFTSINFORMATIK 42(2):165-167.
- [15] Sailer, M (2009): Anforderungsprofile u. akademischer Arbeitsmarkt. Waxmann, Münster.
- [16] Simon, B (2006): Neue Geschäftsmodelle für Bildungsangebote von Hochschulen. Zeitschrift für Betriebswirtschaft Special Issue (2):105-123.
- [17] Simon, B (2010): Gestaltungstheoretische Überlegungen zu Kompetenzmanagementsystemen. WIRTSCHAFTSINFORMATIK 52(6):327-337.
- [18] Strohmeier, S (2008): Informationssysteme im Personalmanagement: Architektur – Funktionalität – Anwendung. 1. Auflage. Vieweg + Teubner, Wiesbaden.
- [19] TIOBE Software (2011): TIOBE Programming Community Index for August 2011. <http://www.tiobe.com/index.php/content/paperinfo/tpci/index.html>. Abgerufen am 4.9.2011.
- [20] Zhu, W-D; Iwai, A; Leyba, T; Magdalen, J; McNeil, K; Nasukawa, T; Patel, N; Sugano, K (2011): IBM Content Analytics Version 2.2. 2. Auflage. IBM, Armonk.